

The Importance of Suspense and Surprise in Entertainment Demand: Evidence from Wimbledon

Paolo Bizzozero*, Raphael Flepp, Egon Franck

November 2, 2015

Abstract

This paper empirically examines how suspense and surprise explain demand for entertainment. We use the Wimbledon Championships tennis tournament as a natural laboratory. This setting allows us to both operationalize suspense and surprise using the audience's belief about the final outcome of the match and observe of the demand for live entertainment using TV audience figures. Using 8,563 minute-by-minute observations from 80 matches between 2009 and 2014, our fixed-effects estimates show that both suspense and surprise are driving forces behind media entertainment demand. Our results have important implications for the design of entertainment content to maximize entertainment demand.

JEL Classification: D83, L82, L83

Keywords: Suspense, Surprise, Entertainment, TV Audience, Betting Odds, Tennis

*Corresponding author. University of Zurich, Department of Business Administration, Affolternstrasse 56, CH-8050 Zurich, Email: paolo.bizzozero@business.uzh.ch

1 Introduction

Media entertainment plays an important role in people’s daily life (Vorderer et al., 2004). Vorderer et al. describe entertainment as enjoyment from consuming media content, whether at home or at an outside venue. Given that entertainment providers face increasing competition in the entertainment market, understanding precisely what factors drive demand for entertainment content is of critical importance.

Previous theoretical studies have identified *suspense* and *surprise* as two major determinants of enjoyment from media consumption (e.g., Zillmann, 1991, 1996; Vorderer et al., 2004). The online Cambridge English Dictionary defines suspense as “a feeling of excitement or anxiety while waiting for something uncertain to happen” and surprise as “an unexpected event, or the feeling caused when something unexpected happens.” Importantly, both occur exclusively in situations of concern over uncertain outcomes (Comisky & Bryant, 1982).

This paper documents field evidence for whether suspense and surprise influence the enjoyment accruing from consuming a media entertainment product. We focus on the relationship between suspense, surprise and the TV audience for 80 men’s singles tennis matches that were transmitted live between 2009 and 2014. Our setting is the Wimbledon Championships tennis tournament, which, worldwide, enjoys a high degree of media coverage with an estimated global TV audience of 1.2 billion (Telegraph, 2015). We present an analysis based on two data sets, each of which demonstrates that both suspense and surprise are important factors in determining the TV audience for tennis matches.

Suspense and surprise can be best understood and modeled in a Bayesian setting (Ely et al., 2015). In this setting, probabilities quantify the personal beliefs: People form hypotheses about the occurrence of specific events (“it will rain tomorrow”) and attach probabilities to them according to their subjective degrees of belief in these hypotheses (“with a 90% probability”). According to Bayes’ rule, upon the arrival of new, relevant information, people will transform their *prior* beliefs into *posterior* beliefs (Itti & Baldi, 2009). In the Bayesian view, the continuous process of forming and updating beliefs generates entertainment through the experience of suspense and surprise, with suspense the future-oriented emotion and surprise the backward-looking one, as we now explain in detail.

Suspense evolves through the assessment of future events, with a moment carrying more suspense when some crucial uncertainty is about to be resolved (Vorderer et al., 2013); for example, a researcher about to open the letter with the committee’s decision on his or her research grant application. These moments are suspenseful because they precede the arrival of new, relevant information that will induce an update in people’s beliefs.

In contrast, surprise arises immediately after an unexpected event occurs. Therefore, because the prior expectations are no longer consistent with the new state of the world, an event causes more surprise when new information has induced a large update in people’s beliefs (Itti & Baldi, 2009). Moments immediately following such events as an underdog baseball team scoring the winning run are surprising because they follow the arrival of new, relevant information that caused a large shift in the individual’s beliefs.

Surprisingly, the empirical evidence on the relationship between suspense and surprise on enjoyment is very limited. Moreover, most of the studies are laboratory experiments. For example, Bryant et al. (1994) recorded a football match and manipulated its commentary to create a high-suspense version and a low-suspense version. Participants, after having watched one of these two versions, were asked to rate their enjoyment on a questionnaire, on a scale from 0 to 10. The results of Bryant et al. (1994) show that viewing the high-suspense version was significantly more enjoyable and exciting. Su-lin et al. (1997) and Peterson and Raney (2008), using students as participants, examine suspense as a factor in the enjoyment of National Collegiate Athletic Association (NCAA) men’s basketball games. Both studies operationalize suspense as the final point differential in a game and enjoyment as the average of seven different items (e.g., “the game excited me” or “I enjoyed the game”), rated on a scale from 0 to 10. Their results show that higher suspense leads to greater enjoyment.

In an experimental setting, Itti and Baldi (2009) test whether surprise attracts the attention of participants when watching one of several videoclips, including television broadcasts such as news, sports, commercials, and outdoor scenes. They define surprise in Bayesian terms as the distance between the prior and posterior distributions of beliefs, and employ eye-tracking technology to measure attention in dynamic natural scenarios. Their results show that surprise explains the greatest fraction of human eye movements, indicating that humans are attracted

towards surprising elements in video displays.¹ Moreover, in a laboratory setting, [Alwitt \(2002\)](#) finds that viewers perceive suspenseful TV commercials as shorter, attributing this effect to the viewers’ intensified attention and interest.

Especially in the sports economics literature, the term “suspense” has been often associated with the concept of “uncertainty of outcome.” The uncertainty of outcome hypothesis states that more uncertainty – about the final winner of a sports competition – causes more suspense ([Borland & MacDonald, 2003](#)). [Alavy, Gaskell, Leach, and Szymanski \(2010\)](#) use minute-by-minute audience figures of 248 English Premier League matches to measure the effect of outcome uncertainty on entertainment demand. They find that matches with less probability of ending in a draw, but also with less score differential between teams, generate more viewers. In a sport-related article, [Olson and Stone \(2014\)](#) model viewers’ entertainment as a function of suspense to evaluate whether the introduction of post-season playoffs in U.S. college football would be an improvement over the current system. Using match-level Nielsen ratings over 2011-2013 for 70 football and basketball matches, they show that the level of viewers’ entertainment significantly increases in the championship’s suspense level.

The formal economics literature on suspense and surprise is also small. Yet [Ely et al. \(2015\)](#) recently filled this gap by introducing a framework in which a Bayesian audience derives entertainment utility (enjoyment) from *anticipated* changes in beliefs (suspense) and *actual* changes in beliefs (surprise). In their model, higher suspense results from greater variance in the next period’s beliefs, i.e., what is currently happening versus what is expected to happen next, and higher surprise results from a greater distance between the previous and the current beliefs.

However, testing this entertainment theory in a natural setting is challenging, because measuring people’s beliefs and entertainment utility in real-world situations is difficult. Therefore, previous studies were conducted mainly in laboratory settings. Nonetheless, despite some clear advantages of rigorous laboratory experiments, testing this theory in the real world may lead to new insights.

Using high-frequency field data, in this paper we test the hypothesis that suspense and surprise, as defined by [Ely et al. \(2015\)](#), drive entertainment demand. Professional tennis offers two unique advantages that help us to tackle the beliefs and entertainment measurement prob-

¹[Baldi and Itti \(2010\)](#) build on these results and provide further evidence of this relationship. For a recent review of studies on the surprise-attention link, see [Horstmann \(2015\)](#).

lems that arise from analyzing real situations. The first advantage is that we can quantify the audience’s beliefs because modeling tennis situations is possible. A Bayesian audience assigns probabilities to a certain hypothesis and updates them as new information is obtained. In tennis, beliefs are formed about the final outcome of the match, i.e., about the likelihood of a player’s winning the match.² As we cannot directly observe individual beliefs, we use two methods to estimate the aggregate audience “belief paths” – this term indicating the evolution of beliefs over time.

The first method, the Markov chain one, relies on the explicit structure of the data-generating process in tennis, particularly on the scoring system. We estimate the unique belief path for each match by computing the likelihood of winning the match for a given player, point-by-point over the match. As a tennis match is composed of points, games, and sets, we model the match as a binary Markov chain (Newton & Aslam, 2009).³ The only input for this simple model is the score and the probability of winning a point on serve.

The second method, the betting odds one, relies on the information content of betting odds from *in-play* betting markets, where betting continues during the event. Importantly, whereas the Markov belief is based only on the actual score and a fixed server’s probability of winning a point, the odds-derived belief is based on all the information acquired second after second by the bettors. These bettors usually follow the event live on TV, continuously betting as the match unfolds. The trading volume on these markets can be very large. For example, \$78.3 million of bets were placed on the 2014 Wimbledon men’s final through Betfair, the online sports betting platform from which we obtain the betting data. Remarkably, 92% of the total betting volume was generated *during* the match. From these in-play betting odds, we can derive the aggregate market’s belief, for all the points in a match, about a given player’s probability of winning the match. Once the beliefs are estimated for every point, we then compute suspense and surprise similarly to Ely et al. (2015) at the minute-by-minute level.

The second advantage that tennis offers is that the demand for entertainment is observed

²In other settings, the same idea can be applied. For example, people assign probabilities to the hypothesis that a president will be reelected, that a mission will succeed, or that a company’s earnings will beat analysts’ consensus estimates. What likely differentiates tennis from other settings is the frequency at which events happen and at which new information is revealed. In our sample, we observe an event average of 1.48 points per minute. In contrast, in football, beliefs are updated when runs are scored, something that may or may not happen during a particular game.

³Appendix A introduces the rules of tennis. Walker et al. (2011, p. 490) illustrate the binary Markov scoring rule for a game in tennis.

through high-frequency minute-by-minute TV audience figures (ratings) during the matches. As viewers can easily, and at no cost, switch channels or turn off the TV to maximize their utility from viewing, short-term variations in TV audience figures reflects whether the audience is enjoying the match (Alavy et al., 2010). Following the idea that the audience has a preference for suspense and surprise, we assume that TV ratings at any minute within the match are stable or increasing for more suspenseful moments, more surprising moments, or both. Additionally, given the increased use of mobile devices with Internet access, viewers might decide to tune in after receiving information about the quality of the match through live feeds, the radio, social networks, or another TV channel’s sports news. By using minute-by-minute information about aggregate viewers’ behavior, measured through audience figures, we are able to uncover the audience’s underlying preferences for entertainment in a real-world environment.⁴

Our empirical analyses show that the effect of both suspense and surprise on entertainment demand is positive. From a fixed-effect regression using 8,563 minute-by-minute observations from 80 matches between 2009 and 2014, we find that minutes with more surprise and suspense show significantly higher live TV ratings. In particular, we estimate that an increase by one unit in suspense based on beliefs derived from the Markov chain model increases the number of TV viewers by 29,000, whereas an increase by one unit in surprise is estimated to increase the number of TV viewers by 8,300 viewers. In contrast, our estimates for suspense and surprise based the betting odds model are 7,100 and 9,700, respectively. In both specifications, suspense and surprise are jointly significant. This result indicates that suspense and surprise are complementary and that demand for entertainment is higher for higher levels of suspense and surprise. Moreover, Su-lin et al. (1997) find that women’s enjoyment drops for extremely suspenseful games. In contrast, our results are equally valid for male and female viewers and are also robust to a disposition towards players (Zillmann et al., 1989) – the enjoyment of watching sports contests may be a function of the viewer’s sentiment toward a particular player.

Our results represent the first time that the theory on Bayesian suspense and surprise has been tested jointly under natural conditions. We provide a framework that entertainment industry managers can use to measure beliefs, and thus entertainment, from suspense and surprise.

⁴TV remains the central provider of entertainment content despite the increasing supply available on the Internet. According to the U.S. Bureau of Labor Statistics, in 2013 people aged 15 and over watched TV for 2.8 hours per day on average, accounting for more than half of their leisure time.

However, for tennis, associations or tournament organizers cannot do much to further increase suspense and surprise, and, in turn, the public’s interest in the event. Practically, although the timing of commercials can be improved, the rules of the game and the skills and effort of the players determine the level of suspense and surprise during the match.⁵ Nonetheless, far more important are the implications of our study for other entertainment settings in which the content can be designed ad hoc for increasing the public’s enjoyment. Designers of films, TV series, TV shows, online videos, novels, or gambling games should be aware of people’s preferences for surprise and suspense.

The remainder of this paper is organized as follows. Section 2 describes our setting and data. Section 3 outlines the operationalization of suspense and surprise and our empirical methodology. Section 4 discusses our main findings. Section 5 concludes and discusses implications for future research.

2 Setting and Data

The Wimbledon Championships tennis tournament, held in London since 1877, is the oldest of the four prestigious major tournaments.⁶ During the match, the arrival of new information continuously affects the audience’s beliefs about a player’s chances of winning. Clearly, a match provides moments of various levels of suspense and surprise: Comebacks, break points, tie-breaks, injuries, or spectacular rallies are making any moment more entertaining, whereas unimportant points are making any moment less entertaining to watch.

We use the 2009 Wimbledon men’s final between Andy Roddick and Roger Federer to illustrate an example. Roddick, after winning the first set, had four set points in the second set, putting him only one set away from the championship.⁷ Federer, however, supported by his strong service, won all of Roddick’s set points and eventually won the set. As the audience

⁵However, it seems that some past rule changes were intended to increase the entertainment. For example, the tie-break was introduced in the 1970s “in an attempt to reduce the chance of spectator boredom with matches which seem to last forever and probably also to make tennis more attractive to the TV networks” (Carter & Crews, 1974). Moreover, in January 2015, Roger Federer played against Lleyton Hewitt in Sydney in a match with new rules (e.g., without advantage scoring and with sets of first-of-four games) aimed at reducing the length of each game and thereby increasing the entertainment. Finally, the actual system of seeding also contributes to ensuring that the best players will meet towards the end of the tournament.

⁶The other Grand Slam tournaments are the Australian, French, and U.S. Opens.

⁷At Wimbledon, the matches are played best of five sets. A player wins the match if he wins a total of three sets (see also Appendix A).

had to strongly readjust their beliefs about Federer’s chances of winning, we describe such a circumstance as surprising. In the final set, which is played until one player wins at least six games with a two-game lead, Federer finally won 16-14 after more than 90 minutes.⁸ Because every point could bring a player’s winning probability either very close to one or very close to zero, we describe such a circumstance as suspenseful.

To quantify such circumstances of high or low suspense and high or low surprise at the minute-by-minute level, we first estimate the relevant beliefs. In tennis, viewers simply form and continuously update their beliefs about the hypothesis of a player’s winning the match. Therefore, we first need to know what happened during the match. Detailed data is provided by IBM.⁹ Beyond general information about the match such as players, courts, start and end match times, this data also contains point-by-point information on the current score, time (exact to the second), server, and winner. We then estimate the winning probability of a player over the length of each match in our sample. To do so, we exploit not only our knowledge of the tennis rules to build our own estimation model but also the information content of the in-play betting odds. As we will explain in greater detail, in the first case we use the server’s probability of winning a single service point, together with the current score, to estimate a player’s likelihood of winning. In the second case, we use the aggregate beliefs from prediction markets.

The in-play betting odds come from Betfair, one of the largest online betting markets. Remarkably, \$1.2 billion bets were placed in 2014 alone, for a total trading value of roughly \$92 billion.¹⁰ Its online platform provides a market for opinions and for participants to bet against one another by offering and accepting odds under which they are willing to buy or sell a certain bet. The odds are formed through a continuous double auction process between buyers and sellers, thus reflecting aggregate bettor beliefs about the outcome of a certain bet (Franck et al., 2011). More generally, Wolfers and Zitzewitz (2006) show that betting odds provide valuable estimates of average, aggregate beliefs about the probability of an event’s occurring.

We use betting odds on the expected match winner, the inverse of which can be interpreted as the aggregate current belief. For each tennis match, we have second-by-second odds information

⁸The classic tie-break rule, where at 6-6 a set is decided by a special tie-break game, does not apply to the final fifth set in matches at Wimbledon.

⁹IBM is the Official Supplier of Information Technology to the Wimbledon Championships.

¹⁰Moreover, between 2005 and 2007, the daily trading intensity on Betfair was larger than that on all the major European Stock Exchanges combined (Croxson & Reade, 2014).

from the in-play betting market, where bettors mostly follow the match live on TV or on other electronic devices, and place their bets during the match.¹¹ Whenever new information becomes observable, bettors update their beliefs, and the odds change accordingly. For the matches in our sample, we observe a considerable average total trading volume of \$28.5 million per match, 70% of which was traded in-play.

To measure entertainment demand, we gather high-frequency TV audience ratings over the 2009-2014 period on all Wimbledon men’s singles matches that were transmitted live and in their entirety on the Swiss national German-language channels *Schweizer Fernsehen Zwei* (SRF2) and *Schweizer Fernsehen Info* (SRFinfo), two free channels of the largest Swiss broadcaster, SRG SSR.¹² In Switzerland, Wimbledon, and tennis as a whole, enjoys a good TV coverage.

The sample is restricted to singles matches transmitted in their entirety, as partially transmitted matches might introduce a sampling bias. For example, the broadcaster may decide to transmit only the final, and presumably most exciting, part of a match, in which case our results would not generalize to *any* moment of *any* match.¹³ Overall, SRF broadcast 108 Wimbledon men’s singles matches between 2009 and 2014; of those, 80 were fully transmitted.

Mediapulse, a Swiss ratings firm essentially equivalent to Nielsen, generates audience statistics using survey data from a panel encompassing 1,870 households across Switzerland, a panel containing approximately 4,200 people aged three years or older. As the advertising industry also uses Mediapulse’s data, the panel has to meet strict requirements so to reflect the Swiss population as accurately and representatively as possible.¹⁴ TV audience ratings measure the total number of single viewers watching the channel at every moment. For example, a rating of 122,500 indicates that 122,500 single viewers were tuned into the program on average during a particular minute. For any minute when more people tune into the tennis match (either from another channel or by turning on the TV) than turn off the match, the rating will increase.

Table 1 reports summary statistics for the dependent variable.

¹¹Fracsoft, an authorized third-party betting data vendor, provides historical betting odds.

¹²While SRF2 focuses on either live or recorded sports programming, SRFinfo chiefly rebroadcasts programs from the SRF1 (the first national channel) and SRF2. However, it also occasionally serves the role of a complementary channel for live sportscasts in case of programming conflicts.

¹³We observe that the start of the match is the part most often cut. Moreover, first-round matches aired from the beginning are more frequently subject to cuts in case of a scheduling conflict.

¹⁴The quota attributes used for determining the appearance of a household in the panel are language area, canton (a member state of Switzerland), district (a region of a canton), household size, presence of children in the household, and age of the head of household. See Appendix B for further details.

Table 1: Summary statistics of TV audience

Variable	Description	N	Mean	Std. Dev.	Min	Max
<i>audience</i>	Rating (in thousand)	8,563	106.47	136.81	0	1,083.7

Notes: Displayed are summary statistics for the 80 men’s singles matches played at Wimbledon between 2009 and 2014, for a total of roughly 8,500 minutes of live tennis.

After we merge the data from the three sources, our final data set consists of roughly 8,500 minute-by-minute observations on detailed match statistics, in-play betting odds, and live TV ratings. Table 2 provides some general information about the matches in our sample. As the table shows, our final sample is very heterogeneous: It includes more than 18,500 points, 58 different players, and matches of different lengths and from different stages of the tournament.

Table 2: Descriptive statistics of the matches

Variable	N	Length (min)	Points (tot.)	Sets	Games
Unique players	58	-	-	-	-
1st stage	12	64	101	2.15	5.59
2nd stage	14	64	102	2.16	5.77
3rd stage	12	82	121	2.40	5.55
4th stage	10	84	118	2.41	5.78
Quarter-final	14	89	128	2.52	6.04
Semifinal	12	91	127	2.47	6.38
Final	6	107	150	2.62	7.00

Notes: The table reports the aggregate summary statistics, obtained from IBM, for the 80 men’s singles matches in the final sample.

3 Estimation Approach

3.1 Operationalization of Suspense and Surprise

The Markov Chain Method

Our definition of suspense and surprise, as previously explained, is best understood in a Bayesian framework. Simply put, a Bayesian audience has some prior beliefs, based on the information currently available, about a certain outcome. Upon the arrival of new, relevant information, the audience will update its beliefs, which are then called posterior beliefs. In the context of tennis,

viewers form beliefs about the hypothesis that a certain player will win the match. For example, the viewers initially assign a probability to the hypothesis “Player 1 will win the match”, and then update this probability over time. To estimate the belief path for each match, we derive, point-by-point, the probability of winning for a given player i and the probability distribution for the following point, by modeling each match as a Markov chain (O’Malley, 2008). In this way, points are linked to games, games to sets, and sets to match outcomes.¹⁵

As the score history of every match is different, we derive unique belief paths for every match. The dotted line in Figure 1 shows a belief path for a particular tennis match in our sample. Our approach to operationalizing suspense and surprise closely follows the work of Ely et al. (2015), in which the entertainment utility of the Bayesian audience is a function of the belief path. First, we model suspense in the form of an expectation, where higher suspense results from greater variance of the next period’s belief. For the point p :

$$SUS_p^{markov} = [E \sum_{\omega} (\tilde{\mu}_{p+1}^{\omega} - \mu_p)^2]^{1/2}, \quad (1)$$

where μ_p refers to the probability of a player’s winning the match (the current belief) at the moment when point p was scored and $\tilde{\mu}_{p+1}^{\omega}$ to the anticipated posterior probability of a player’s winning the match (the posterior belief). By construction, the expectation of the posterior is the prior ($E[\tilde{\mu}_{p+1}] = \mu_p$) and μ_p is always equal to 0.5 at the beginning of the match. The posterior belief depends on the realization of the state ω , which in this setting is a binary variable: $\omega = 1$ when the server wins the point (with probability S), or $\omega = 0$ when he loses it (with probability $1 - S$). We assume that $S = 0.66$, which is the historical average probability of serve points won at Wimbledon.¹⁶ As in reality only one state occurs, we also estimate the *counterfactual* belief,

¹⁵As in Ely et al. (2015), we code the model using Python, a widely used programming language, following the procedure developed by Jeff Sackmann. Sackmann’s original Python code is freely available at www.heavytopspin.com. The studies of Liu (2001) and Barnett and Clarke (2005) propose similar ways of modeling the probability of winning a match. As noted previously, Walker et al. (2011, p. 490) illustrate the binary Markov scoring rule for a game in tennis that underlies this model.

¹⁶Wimbledon is the only Grand Slam tournament played on grass, which is known for faster ball bounces than clay and hard courts. Thus, that dominant servers are favored is reflected in a higher average winning percentage for serve points. For the 2007 Grand Slam events, Newton and Aslam (2009) collect field data on the probability of winning on serve: 65.99% for Wimbledon, 63.67% for the U.S. Open (hard), 62.36% for the Australian Open (hard), and 61.67% for the French Open (clay). The probability of winning the point on serve S could be measured differently. For example, it could be computed for each player individually from historical data. This alternative, which we implement in the section with betting odds, makes very similar predictions as assuming $S = 0.66$.

defined as the probability of winning the match for the unobserved state. To do so, we replace the actual score with the counterfactual score in the model. For example, if player i actually wins the first point of the match and leads 15-0, the counterfactual belief is computed by assuming that he *lost* the first point and, because he would lose 0-15, his probability of winning would be less than 50%.

Second, we model surprise in the form of the Euclidean distance between the last period’s belief and the current belief, where higher surprise results from the occurrence of an event that strongly contradicts the audience’s belief, constraining the viewers to change their beliefs (Itti & Baldi, 2009). For the point p :

$$SUR_p^{markov} = |\mu_p - \mu_{p-1}|, \quad (2)$$

where μ_{p-1} refers to the probability of winning the match for a player one point earlier. In Appendix C, we use an example to illustrate how we compute SUS^{markov} and SUR^{markov} .

Because we observe variation in the TV audience at the minute level, we translate the suspense and surprise from point level to minute level by computing the average suspense or surprise over the minute when more than one point is scored within a minute.¹⁷ Table 3 reports summary statistics for the suspense and surprise variables, based on the Markov method.

Table 3: Summary statistics of SUS^{markov} and SUR^{markov}

Variable	Description	N	Mean	Std. Dev.	Min	Max
SUS^{markov}	Suspense based on the Markov model	8,563	0.0196	0.0190	0	0.2734
SUR^{markov}	Surprise based on the Markov model	8,563	0.0181	0.0216	0	0.9154

Notes: Displayed are summary statistics for the 80 men’s singles matches played at Wimbledon between 2009 and 2014, for roughly 8,500 minutes of live tennis.

¹⁷We also test an alternative method, which considers only the last point scored in any minute when more than one point is scored. This change does not alter our results.

The Betting Odds Method

The construction of suspense and surprise based on betting odds is straightforward. Following Hasbrouck (1991), we compute the average mid-odds from the best buy $odds_{ip}^{back}$ and sell $odds_{ip}^{lay}$ for the player i for each point p in the match m as:

$$odds_{ip}^{mid} = \frac{[odds_{ip}^{back} + odds_{ip}^{lay}]}{2} \quad \forall(p, m), \quad (3)$$

from which we derive the implied winning probability for player i :

$$\tilde{\nu}_{ip} = \frac{1}{odds_{ip}^{mid}} \quad \forall(p, m). \quad (4)$$

Although the sum of the winning market probability for player i and player j ($\tilde{\nu}_{ip} + \tilde{\nu}_{jp}$) should sum up to one in a frictionless market, in practice it rarely does so, due to transaction costs. Following the standard approach to eliminating this overround (e.g., Forrest et al., 2005), we adjust the implied winning probability to obtain the final market implied winning probability for player i as follows:

$$\nu_{ip} = \frac{\tilde{\nu}_{ip}}{\tilde{\nu}_{ip} + \tilde{\nu}_{jp}} \quad \forall(p, m). \quad (5)$$

The solid line in Figure 1 illustrates a belief path computed using the betting odds for a tennis match in our sample. In general, the higher a belief's volatility, the larger the suspense and surprise. Moreover, it can be noted that the belief path from the Markov model correlates strongly with the belief path extracted from the betting odds.¹⁸

We then define surprise for the point p as:

$$SUR_p^{odds} = |\nu_p - \nu_{p-1}|, \quad (6)$$

where ν_p refers to the probability of a player of winning the match (the current market belief) at the second point p was scored. As the change in the implied winning probability is symmetric,

¹⁸Recall that the only inputs for the model are the current score and the probability of winning the point on serve. Apparently, bettors also strongly rely on this information. The belief paths might diverge for the first moments of a match (as μ_p is always equal to 0.5 at match start) or when new, relevant information (beyond server and score) is revealed to the audience, e.g., a slight injury.

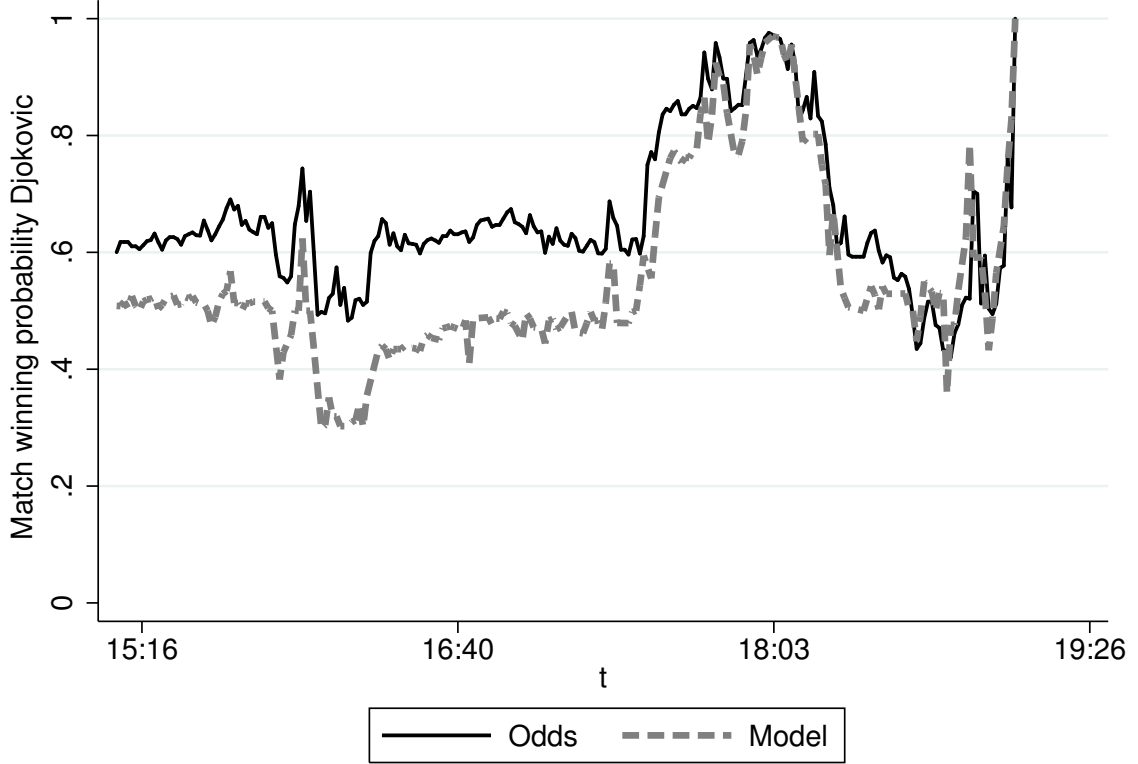


Figure 1: Displayed are the belief paths about Novak Djokovic’s chance to win the 06/07/2014 match played against Roger Federer. Djokovic won in five sets (6-7⁷⁻⁹; 6-4; 7-6⁷⁻⁴; 5-7; 6-4 – the superscripts indicate the result of the tie-break). The Markov model-generated beliefs (indicated by the dotted line) start by definition at 50% and end at 0% for the player who lost, in this case Federer. The betting odds are from Betfair: a total volume of \$78.3 million was bet on this final.

the choice of which player’s probability of winning to consider is irrelevant. Thus we can drop the player subscript i . The only difference from Equation (2) lies in the type of data used.

Second, as we cannot implement the baseline metric of suspense by relying entirely on betting odds, we slightly redefine the baseline suspense measure. By definition, betting odds reflect the bettor’s current beliefs, based on a given information set. Thus we do not know which value the odds would have taken had the point gone differently, i.e., the counterfactual. Therefore, we come up with a “hybrid” implementation of the baseline suspense measure, where we partially use the Markov chain model to determine the change in beliefs for the counterfactual. As the audience also has to estimate the counterfactual probability, this procedure is suitable for our purposes. We model suspense for the point p as follows:

$$SUS_p^{odds} = [S_i \cdot (\nu_{p+1}^\omega - \nu_p^\omega)^2 + (1 - S_i) \cdot (E[\tilde{\mu}_{p+1}^\omega] - \tilde{\mu}_p^\omega)^2]^{1/2}, \quad (7)$$

where $E[\tilde{\mu}_{p+1}]$ is the counterfactual probability and $\tilde{\mu}_p$ is the previous point counterfactual probability computed using the Markov chain model. The major difference between Equations (1) and (7) is that we substitute μ_p with the implied probabilities from the betting odds ν_p for the prior and observed posterior beliefs. Moreover, we newly compute the entire posterior distribution using a player-and-year-specific win probability on serve S_i . The advantage of this procedure is that we use the same statistics that are also readily available to bettors.¹⁹

As before, we translate the suspense and surprise from point level to minute level, using the average suspense or surprise over the minute when more than one point is scored within a minute. Table 4 reports summary statistics for the suspense and surprise variables based on the in-play betting odds. A comparison with Table 3 shows that the mean of SUS^{odds} is slightly higher

Table 4: Descriptive Statistics of SUS^{odds} and SUR^{odds}

Variable	Description	N	Mean	Std. Dev.	Min	Max
SUS^{odds}	Suspense based on the odds	8,563	0.0233	0.0256	0	0.2601
SUR^{odds}	Surprise based on the odds	8,563	0.0129	0.0197	0	0.3097

Notes: Displayed are summary statistics for the 80 men’s singles matches played at Wimbledon between 2009 and 2014, for a total of roughly 8,500 minutes of live tennis.

than the mean of SUS^{markov} , whereas the mean of SUR^{odds} and SUR^{markov} are comparable. This result is reasonable: Although we are trying to measure the same phenomenon, namely the level of suspense and surprise at any minute, we are using two different data sets. Importantly, the model belief is based only on the actual score and server, whereas the odds belief is based on all the information acquired by the bettors, e.g., a player’s minor injury.

3.2 Empirical Methodology

The type of data used in this study presents two advantages with respect to the estimation methods. First, our panel data allows us to control for time-invariant factors that might jointly affect the audience level by using within-match variation. Those factors could be the stage of

¹⁹We download the data from www.atpworldtour.com, where official player-level statistics are available. For all the 160 player-year combinations in our sample, we download the player’s average “1st serve in”, “1st serve points won”, and “2nd serve points won” statistics from the previous year on grass courts. From that, we compute each player’s yearly statistics “points won on service” S_i using the formula: $S_i = [(1^{st} \text{ serve in}) * (1^{st} \text{ serve points won}) + (1 - 1^{st} \text{ serve in}) * (2^{nd} \text{ serve points won})]$. For players with no history on grass courts for the previous year, we use statistics from two years earlier or, when also unavailable, from the previous year but on hard courts. Federer has the highest S of our sample (78.7% in 2009), whereas Albert Ramos-Vinolas has the lowest (48.3% in 2012).

the competition, the day of the week, the quality of the players or of the match, or the audience level of the previous TV program.²⁰ Second, as opposed to stadium attendance, short-term TV audience variation is not affected by factors such as supply capacity, gate price, location, officials, audience size, quality of viewing, or weather conditions (Borland & MacDonald, 2003; Alavy et al., 2010).

At the same time, to account for other characteristics that are subject to change *during* a match and that might affect the minute-by-minute TV ratings, we introduce several control variables. In the evening hours, all channels offer their best TV content, which may turn viewers away from tennis. Thus we construct an indicator variable, *primetime*, that takes the value of one for all the minutes after 6:00 p.m. and zero otherwise. Intuitively, the coefficient for *primetime* should be negative.

Due to the popular news program on SRF1, the first national channel, the TV audience level of SRF2 might drop as viewers switch to the newscast. Therefore, we construct the *news* indicator variable, which equals one for any minute between 5:58 - 6:06 p.m. and between 7:28 - 7:56 p.m., indicating the first short newscast (6:00-6:05 p.m.) and the following long one (7:30-6:55 p.m.), respectively, and zero otherwise. Overall, anticipating the sign of the *news* coefficient is difficult, as it should capture the effect of viewers switching channels for the news but also their returning to the tennis match. Other programs are not likely to systematically affect the audience variation, especially because the matches take place at various times of the day, from Monday through Sunday. As Alavy et al. (2010) note, minor variations in audience might be due to channel hoppers. Nonetheless, they argue that any moment providing high entertainment should attract even channel hoppers and keep them watching, thus reducing the noise from their behavior.

To allow the players to rest and switch sides of the court, small breaks take place after odd-numbered games and between sets. As these breaks may cause viewers to briefly stop watching, we introduce the *pause* indicator variable, which equals one during the break and zero

²⁰Rodríguez et al. (2015) illustrates the importance of including a large set of control variables when examining aggregate TV audience measures. As their dependent variable is the average TV audience over the length of the program (they analyze professional cycling races), they also control for, among other things, calendar variables, audience inertia, the scheduling of rival channels, and the competitive balance before the race.

otherwise.²¹ The *pause* variable is needed only for the regression based on betting odds. By construction, the suspense and surprise measures computed with the Markov model are missing during the break, because the score does not change.

Therefore, the general form of the regression is:

$$\begin{aligned} audience_{i,t} = & \beta_0 + \beta_1 \cdot SUS_{i,t} + \beta_2 \cdot SUR_{i,t} \\ & + \beta_3 \cdot primetime_{i,t} + \beta_4 \cdot news_{i,t} + \beta_5 \cdot pause_{i,t} \\ & + v_i + u_{i,t}, \end{aligned} \tag{8}$$

where the subscripts i and t denote match and minute, and v_i denote the match fixed effects. The dependent variable $audience_{i,t}$ represents the match i 's average live TV audience level at minute t . Moreover, as the Wooldridge test evidenced first-order autocorrelation in the audience measures, we allow the error term to be first-order autoregressive ($u_{i,t} = \rho \cdot u_{i,t-1} + \eta_{i,t}$, where $|\rho| < 1$ and $\eta_{i,t}$ is independent and identically distributed with mean 0 and variance σ_η^2). Diagnostic plots show that the model assumptions are valid.

4 Empirical Results

4.1 Univariate Evidence

Before we turn to the estimation results of Equation (8), we report the results of a univariate analysis of our data. Due to programming conflicts, the broadcasting of tennis is sometimes switched from SRFinfo to SRF2 or vice versa. These switches occur independently of the current match standing and thus independently of the level of suspense and surprise. Figure 2 depicts such an example from a 2012 third-round match between Roger Federer and Julien Benneteau, initially broadcast on SRFinfo and then switched to SRF2 at 9:35 p.m. Clearly, the audience actively switched channels to follow the match, indicating that the audience was actively following.

However, the relevant question is whether the size of the audience change depends on the

²¹TV broadcasters also use some breaks for showing commercials. Therefore, we do not insert an *advertising* dummy because the *pause* dummy already captures the effect of viewers switching channels to skip the advertising. We examine the channel content description for SRF2 and SRFinfo, which lists all the programs broadcast and their exact start and end times, to identify time-outs for advertising. Commercial breaks occur on an average 35-minute basis for all non-final matches and on a 20-minute basis for a final match. As SRF2 is a national public channel, the amount of advertising is limited by law.

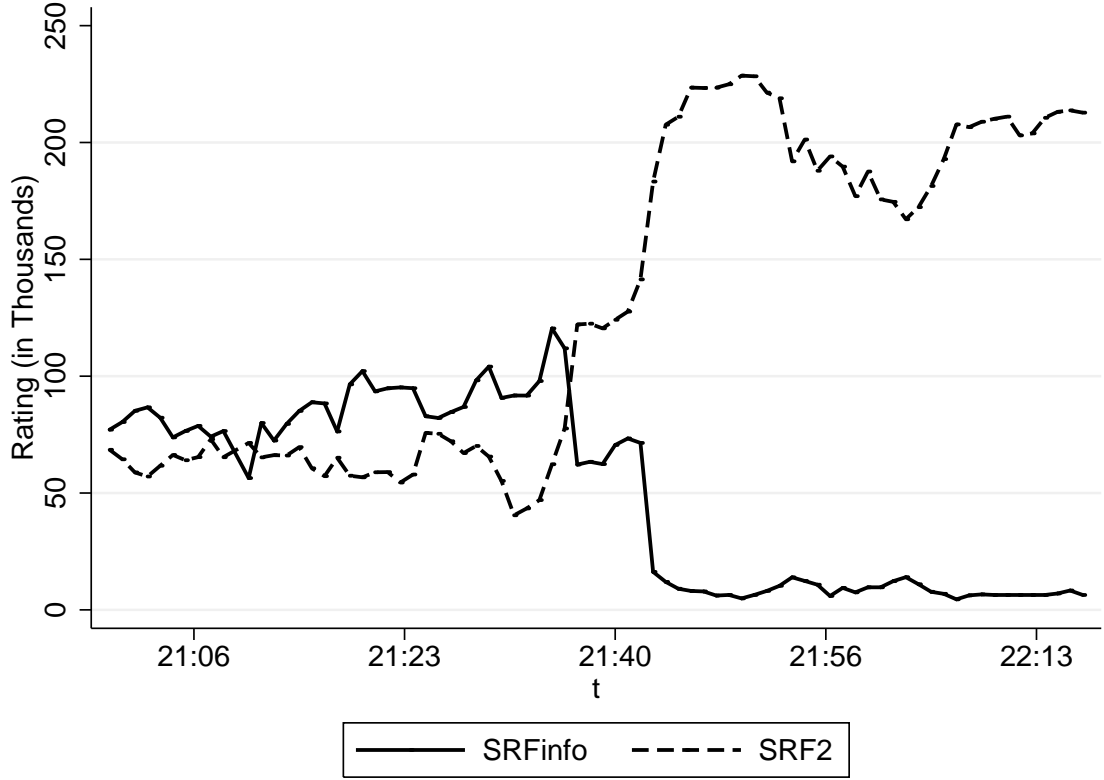


Figure 2: TV audience (rating) evolution around an exogenous channel switch. Both channels broadcast the same content for about three minutes from 9:35 p.m.

level of suspense and surprise in the previous minutes. To answer this question, we identify eight switches that occurred during the broadcast of matches in our sample. For each switch, we can define the precise moment when the live broadcast was interrupted and continued on the other channel. We measure the jump in ratings on the channel where the broadcast is continued by taking the difference of the average 15-minutes ratings before and after the switch. SUS^{markov} and SUR^{markov} are measured as the sum over 15 minutes before the switch. We create two groups based on whether suspense and surprise were above or below the subsample average. Intuitively, if the 15 minutes before the switch offered above-average (below-average) entertainment, this should be reflected in a larger (smaller) jump in the TV audience on the post-switch channel.

We use the Wilcoxon rank sum test to compare differences in TV audience variation after a broadcaster-initiated channel switch between groups with low and high suspense and surprise. We find that the group with below-average suspense and surprise shows an increase in TV

audience of 29,900 units, while the group with above-average suspense and surprise shows an increase in TV audience of 99,450 units ($z = 2.309$, $p < 0.05$). Overall, these findings not only provide evidence for an active audience assumption, but also provide preliminary, suggestive evidence supporting the hypothesis that a TV program’s level of entertainment, measured in terms of suspense and surprise, impacts its success in terms of audience.

4.2 Regression Analysis

Table 5 reports regression estimates for the effect of suspense and surprise on TV audience ratings. Overall, the results are in line with the hypothesis that moments that offer more suspense and surprise generate more entertainment demand, as reflected in the higher TV audience figures.

Columns (1)-(3) show the results for suspense and surprise measures based on the Markov chain model. Columns (1) and (2) demonstrate that suspense and surprise individually have a positive effect on TV ratings. Even when we include both measures in column (3), suspense and surprise remain significant predictors for the size of the TV audience. This result indicates that suspense and surprise are both driving forces behind media entertainment demand. Moreover, we find that the effects of suspense and surprise on the TV audience are large: According to our estimates in column (3), the viewing audience increases by about 29,600 viewers for each additional unit of suspense while viewing audience is larger by about 8,300 viewers for each additional unit of surprise.

Columns (4)-(6) report the results for suspense and surprise measures based on the in-play betting odds. In all specifications, suspense and surprise have a positive and significant coefficient. Column (6) shows that the viewing audience increases by approximately 9,800 viewers for each additional unit of surprise, a result similar to that in column (3). However, the magnitude of the suspense coefficient is lower than in the results in column (3), as the the audience increases by 7,200 viewers for each additional unit of suspense.

The coefficients of the control variables mostly have the expected sign. We find a negative coefficient for *primetime*, probably indicating an increase in the competitive mix of TV programs offered during the evening. The coefficient for *news* is negative and significant at the 5% level in column (3), whereas it is not significantly different from zero in column (6), thus partly confirming our hypothesis that the daily news on the SRF1 may attract some viewers away

Table 5: The effect of suspense and surprise on TV audience.

Variables	Markov model			In-play betting odds		
	(1)	(2)	(3)	(4)	(5)	(6)
SUS	28.565*** (4.909)		29.61*** (5.543)	7.135** (3.438)		7.165** (3.437)
SUR		12.857*** (3.326)	8.308** (3.527)		9.762** (4.093)	9.793** (4.092)
primetime	-5.928*** (1.430)	-5.731*** (1.412)	-7.073*** (1.488)	-5.565*** (1.527)	-5.566*** (1.527)	-5.564*** (1.527)
news	-1.262* (0.780)	-1.292* (0.785)	-1.488** (.797)	0.070 (0.869)	0.089 (0.869)	0.083 (0.868)
pause				-2.044*** (0.234)	-1.997*** (0.232)	-2.072*** (0.234)
Constant	218.6*** (0.073)	224.2*** (0.072)	242.4*** (0.075)	174.2*** (0.081)	174.2*** (0.082)	174.0*** (0.081)
Match Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	8,448	8,448	8,448	8,483	8,483	8,483

Notes: The table reports results of panel regressions of suspense and surprise derived from the Markov model and in-play betting odds on the TV audience rating. The data is at the second-level and includes 80 men’s singles matches played at Wimbledon that were broadcast live on SRF2 and SRFinfo from 2009 to 2014. In all models, *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

from the tennis match. The coefficient for *pause* is significantly negative, indicating that some viewers rapidly switch channels during the short pauses after odd-numbered games and during the longer breaks between the sets.

4.3 Extension and Robustness Checks

In an experimental study on the effects of suspense on enjoyment, [Peterson and Raney \(2008\)](#) note that the utility from watching a sportscast also depends on a viewer’s disposition towards the participants. The disposition theory describes this effect in detail. It proposes that “enjoyment derived from witnessing the success and victory of a competing party increases with positive sentiments and decreases with negative sentiments towards that party” ([Zillmann et al., 1989](#), p. 162). As our data comes from Swiss households and as 30 matches in the sample include Roger Federer, a Swiss player, the audience might show an affective disposition towards him. Moreover, as Federer has been very successful in the Wimbledon Championships tournament – he ties with Pete Sampras for the most men’s singles championships won (seven) – the perceived entertainment from watching might be positively amplified when Federer is winning or reduced

when he is losing. To deal with this possibility, we perform another analysis in TV audience variation for only the subsample of the matches showing Federer. Our results remain qualitatively the same for this subsample (untabulated results).

As described earlier, [Su-lin et al. \(1997\)](#) investigate spectators' enjoyment at NCAA men's basketball matches. They show that extremely suspenseful matches cause a decrease in enjoyment among some viewers, particularly women. According to [Zillmann \(1996\)](#), suspense also contains aspects of anxiety-like uncertainty, especially if the stakes are too high ([Ely et al., 2015](#)). Thus, extreme levels of suspense might be undesirable, implying a non-monotonic preference for suspense.²² To allow for a non-monotonic relation between suspense and TV ratings, in an additional analysis, we introduce a quadratic term for suspense. We observe that this term is negative, as expected, but not significant at any reasonable significance level. Thus we cannot confirm the results of [Su-lin et al. \(1997\)](#).

Furthermore, as we can split the audience by gender, we test our main regression specification on subsamples of only male viewers and only female viewers. The coefficients remain highly significant and positive for both samples, whereas the size of suspense is slightly smaller in the Markov model regression than in column (3) in Table 5. This evidence suggests that our results are equally valid for male and female viewers.

5 Concluding Remarks

Media entertainment has become central in several domains of people's life. The theoretical literature has suggested that suspense and surprise should be important determinants of enjoyment, with some empirical evidence from laboratory experiments. Nonetheless, an important challenge still facing entertainment theory is that of validating predictions with empirical evidence from natural settings. In this paper, we present the suspense and surprise hypothesis in the real-world setting of the Wimbledon Championships tennis tournament. This unusual panel data set provides a unique opportunity for determining how suspense and surprise explain enjoyment and demand for entertainment. From our minute-by-minute analysis of 80 tennis matches, we

²²Given psychological theories, the same cannot be said for surprise. Even though in real-life situations too much surprise might be unwelcome, in sports high surprise usually implies more entertainment utility, i.e., some interesting, crucial action taking place.

show that both suspense and surprise jointly increase entertainment demand as measured by TV ratings.

In general, as suspense and surprise are important, entertainment industry leaders should take both into consideration. The U.S. Major League Soccer, for example, is a closed league that strongly focuses on high suspense by inducing a fixed number of teams to compete with comparable levels of talent, enforced through uniform salary caps and extensive revenue redistribution. In contrast, in the various European soccer leagues, the absence of salary caps and the low levels of revenue redistribution allow teams to be more heterogeneous in expenditures on talent and consequently in playing strength. This European setting increases the potential for surprise, as underdogs regularly encounter clear favorites and occasionally beat them.

However, the European model does not ignore the value of suspense: relegation and promotion of clubs based on sporting merit in the entire pyramid of interconnected leagues makes sure that disparities in playing strength cannot exceed certain levels within one league. Parallel European club competitions such as the Union of European Football Associations (UEFA) Champions League add additional suspense when they match comparably strong top clubs from all European countries. This anecdotal evidence from soccer leagues suggests that even the very formats of sports competition could be redesigned for improving overall entertainment and success. Furthermore, in the context of the design of sports rules, sport leaders could use our framework to analyze past decisions such as, for example, the introduction of the “golden goal” in soccer, from the perspective of suspense and surprise.

Most importantly, our findings have implications for the design of entertainment content, particularly for products such as movies, TV series, talent contests, or TV shows, where suspense and surprise can be endogenously determined through a rigorous design of their contents. We show that the audience has preferences for being held in suspense and surprised. Our methodology can be applied for effectively measuring how entertaining an audience perceives the product to be. The challenge remains in measuring people’s beliefs. Nonetheless, even for settings where prediction markets are unavailable or where theoretical modeling is extremely problematic, the measurement of beliefs can be feasible.

Social media analytic tools, for example, can be used for analyzing posts and user comments from social media platforms such as Twitter or Facebook. Just as tennis fans talk about and

want to hear about Wimbledon on Twitter, TV series fans will do the same. Alternatively, beliefs could be derived from historical data. For example, to measure the average audience’s belief that a particular movie will have a happy ending, one could use as prior belief the historical fraction of movies of a certain genre with a happy ending and compute suspense and surprise using the actual type of movie end (i.e., the posterior belief). Ideally, detailed data at the individual level could be also used for thoroughly examining individual reactions to various levels of suspense and surprise and customizing a product as a function of them.

The timing of commercials can also be improved. As suspense and surprise generate entertainment and simultaneously attract human attention, commercials will be more effective following very suspenseful or surprising moments. A larger and more attentive audience could be reached, thereby increasing the effectiveness of the advertisement. The pricing of advertising could also be determined as a function of the level of suspense or surprise.

We recognize that the issue of suspense and surprise is multifaceted. Specific time patterns or combinations of suspense and surprise may generate different levels of entertainment. For example, if a moment has just delivered a lot of entertainment, the audience will likely better “tolerate” a slightly boring moment immediately afterwards. Too much suspense and surprise over a long period might be stressful for certain audiences. Finally, our data does not allow us to distinguish the effect of suspense and surprise from other factors that may also determine entertainment demand. Mood management, escapism, or learning motivations, for example, might become relevant in contexts other than sports. Thus assessing the transferability of our findings to other settings will require further research.

Acknowledgements

The authors are grateful to the All England Lawn Tennis Club, IBM, and the *Schweizer Radio und Fernsehen* for providing the data. We are indebted to Alex Frankel for useful consultation. The authors thank seminar and conference participants at the University of Zurich for very helpful comments and suggestions, and Fabio Banfi for helping with Python. We take responsibility for all remaining errors.

Appendix A

In this appendix, we introduce the basic rules of tennis. Player 1 begins the match by serving in the first game of the first set. Player 1 wins a point (sometimes referred as “point game”) if player 2 cannot return the ball. A *game* is won when one player wins four points with a two-point difference, or when there is a two-point difference after a deuce (i.e., a score of 40-40 in a game). The players alternate serving after every game and they change ends after every odd-numbered game. A *set* is won when either a player wins six games with a two game difference, or, in the case of a tie-break when the score for one player is 7:6. The *tie-break* starts when the game score is tied at 6:6, and is played until one player wins seven points with a two-point difference, or until there is a two-point difference when the point score is 6-6. At Wimbledon, a tennis match is played best of five sets (instead of three), meaning that a player winning three sets wins the match. The fifth set does not have a tie-break; The set is won when one player has won two games more than the other.²³

²³For further details on the rules of tennis, see: <http://www.itftennis.com/media/107013/107013.pdf>.

Appendix B

This appendix provides details on the Mediapulse TV panel.

- The TV audience measurement system: Every household in the panel is given a small measuring device which is connected to all TV sets in the house or apartment. This field-tested device is used in almost 25 countries worldwide. It collects audience information every second. Mediapulse then aggregates this data for the entire panel and saves it at the minute-by-minute level. For example, the TV audience corresponding to the minute 15:23:00 indicates the average audience between 15:23:00 and 15:23:59.
- Extrapolation: The reporting samples of a panel are extrapolated to the population (universe) estimates, allowing the results to be representative of the respective population, i.e., of the TV panel target audience (Kuonen & Hulliger, 2013). For extrapolating, Mediapulse uses actual population figures from the Swiss Federal Statistical Office. To certify that the panel conforms to international quality standards, it is subject to an annual external verification.
- Changes in the panel: Over the period 2009-2014, two changes in the measuring system have taken place. First, starting in 2010, daily weighting was introduced, and the use of Teletext was considered to be TV watching. Second, before 2013, the panel of 1,918 households was randomly recruited by the phone. This universe comprised all households with at least an installed TV. However, after January 1, 2013, the panel was recruited by either the phone or mail. It now contains 2,000 households, at least 1,870 of which must provide data daily. For both panels, households watching TV exclusively from a computer are not included. Mediapulse informed us that for an audience analysis *within* a program, like a tennis match, none of these changes has had an impact on the size of the variation in the measured audiences.

Appendix C

Using a fictitious example, we illustrate how to compute SUS^{markov} and SUR^{markov} , i.e., suspense and surprise based on a binary Markov chain. An almost identical procedure applies for computing SUS^{odds} and SUR^{odds} . Player W is playing against player L. After the 153rd point, the match is tied at one set apiece, with player W ahead five games to four (5-4) in the set, and 40 point games to zero (40-0) in the game. This score means that player W is serving for winning the game and, by doing so, also the set (the expression used in tennis is: “Player W has three set points”). We assume that, as the server, player W’s probability of winning next point is 66%. As the situation of the players in a tennis match is always symmetrical, this symmetry implies that player L’s probability of winning on a return point is 34%.

Using this information, the model predicts that player W’s match winning probability is 76.24% (i.e., the actual belief μ_p). If he wins the point, the posterior belief would rise to 80.50% (+4.26%). However, as there is a 34% chance that he will lose next point, that loss would bring the posterior belief down to 75.56% (-0.68%). The posterior belief distribution, for both state of the world ω – the server wins ($\omega = 1$) or loses ($\omega = 0$) – is as follows:

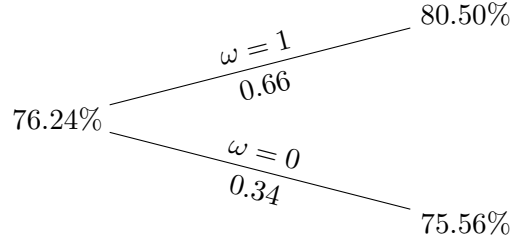


Figure 3: Probability transition from point 154 to 155.

The possible size of the update in the beliefs is clearly asymmetric (+4.26% vs. -0.68%), as player W would still have two set points left to play even if he loses this serve point. Finally, if player W indeed loses the 154th point ($\omega = 0$), the posterior belief would drop to 75.56%, making 80.50% the counterfactual probability. Using Equation (1), we compute suspense for the 153rd point ($p = 153$) as the standard deviation of the next period’s beliefs:

$$SUS_{153}^{markov} = (0.66 \cdot (0.8050 - 0.7624)^2 + 0.34 \cdot (0.7556 - 0.7624)^2)^{1/2} = 0.0348.$$

Similarly, using Equation (2) we compute surprise for the 154th point ($p = 154$) as the absolute value of change in beliefs from period 154 to 155:

$$SUR_{154}^{markov} = |0.7556 - 0.7624| = 0.0068.$$

Thus the model predicts a suspense level of 0.0348 and a surprise level of 0.0068. Note that, as one of the input parameters, the player's probability of winning a service point is held constant throughout the match, the implicit assumption is that points are independent and identically distributed random variables. (Newton & Aslam, 2009) show that this assumption is valid in tennis.